American Journal of Medical Genetics 62:61–67 (1996)

Epidemiological Analysis of the Association Between Two Congenital Anomalies in the Same Child: A Method for Adjusting Nonspecific Clustering

Luis Prieto and María-Luisa Martínez-Frías

Department of Biostatistics (L.P.), ECEMC and Department of Pharmacology (M.-L.M.-F.), Faculty of Medicine, Universidad Complutense, Madrid, Spain

Most of the methods proposed to analyze patterns of congenital anomaly clusters have been criticized because large observed/expected (O/E) ratios are obtained for many defect combinations, suggesting that the association of congenital defects is a generalized and nonspecific tendency. To avoid this problem, two methods have been proposed: (1) the analysis of the association of two defects in the subgroup of infants with only multiple congenital anomaly (MCA) patterns, and (2) a method for adjusting the O/E ratio for the nonspecific tendency of defects to cluster among themselves. However, neither of these methods analyzes the association of *only two* particular defects in a child *who apparently has no other anomalies* apart from the two that are being analyzed.

Here we present a method that analyzes the association of two particular defects, when they are the only defects observed in the child. The method is tested on the Spanish Collaborative Study of Congenital Malformations (ECEMC) data base. The proposed method also controls for the tendency of nonspecific defect clustering.

KEY WORDS: epidemiologic methods, cluster of defects, associations, congenital anomalies

INTRODUCTION

One of the main interests in the study of infants with multiple congenital anomaly (MCA) patterns is to elucidate whether the association of *only two* defects in a child who has no other anomalies, is the result of (1) a chance occurrence, (2) a nonspecific and generalized tendency for both defects to cluster (the same as the tendency for all other defects to cluster among themselves), (3) a nonrandom and specific association, or (4) a pleiotropic manifestation of a specific causal entity, such as a mendelian mutation.

Different methods for analyzing patterns of congenital anomaly clusters have been proposed [Roberts and Powell, 1972; Källén, 1988; Khoury et al., 1990], but most of them have been criticized because large O/E ratios are obtained for many defect combinations, suggesting a generalized and nonspecific tendency for the association of congenital defects. To avoid this problem, Källén [1988] proposed examining the association of two defects in the subgroup of infants with only MCA patterns. Nevertheless, as pointed out by Khoury et al. [1990], this method is affected by the tendency of defects to cluster among themselves which results in spuriously low O/E ratios. These last authors proposed a method for adjusting the O/E ratio for this tendency. However, none of the above-mentioned methods analyzes the association of *two* particular defects in a child *who has no other anomalies* apart from the two that are apparent.

Here we present a method to analyze the association of two particular defects, when they are the only defects observed in the child, as well as when they also exist in combination with other anomalies. The method also controls for the tendency of nonspecific defect clustering. The method is applied to data from the Spanish Collaborative Study of Congenital Malformations (ECEMC).

METHODS

The analysis of a possible preferential association between two particular defects A and B is based on the comparison of the proportion of cases with each one of the two defects in relation to those with or without the other one. That is, the proportion of cases with defect A among the cases *with* defect B, and the proportion of cases with defect A among the cases *without* defect B and vice versa.

Received for publication April 17, 1995; revision received September 21, 1995.

Address reprint requests to Dr. M.L. Martínez-Frías, ECEMC, Facultad de Medicina, Universidad Complutense, 28040, Madrid, Spain.

Considering two defects A and B occurring in proportions P(A) and P(B), respectively, to control for the nonspecific tendency of defects to cluster, we should also take into consideration the presence or absence of congenital defects other than A and B, considered globally as a third factor C. Thus, the proportion of cases with defect A in children with and without defect B and with/without other defects (C) will be estimated. This requires a table that distributes the cases, ignoring the presence or absence of the defect being analyzed (defect A in this example).

The information necessary for the analysis is:

A_1 = Number of infants with isolated defect A.
B_1 = Number of infants with isolated defect B.
O_1 = Number of infants with only defects A plus B.
A_2 = Number of infants with defect A in infants with other congenital defects.
B_2 = Number of infants with defect B in infants with other congenital defects.
O_2 = Number of infants with defects A plus B in infants with other congenital defects.
A = Total number of infants with defect A ($A_1 + A_2$).
B = Total number of infants with defect B ($B_1 + B_2$).
O = Total number of infants with defect A + B ($O_1 + O_2$).
C = Number of infants with other congenital defects.
D = Total number of malformed infants.
T = Total number of births (if this information is available).

The table would be as follows:

	Infants with B	Infants without B
Infants with C	$a = (B_2 - O_1)$ $[a'] = O_2$	$b = (D - B - A_1)$ $[b'] = A_2 - O$
Infants without C	$c = (B_1 + O_1)$ $[c'] = O_1$	$d = (T - D + A_1)$ $[d'] = A_1$
Total	$a + c = (B)$ $[a' + c' = O]$	$b + d = (D - B - A_1) + (T - D + A_1)$ $[b' + d' = (A_2 - O) + A_1]$

where

Infants with C = Children with other defects (excluding infants with only the defect A as an isolated anomaly).

Infants without C = Children without defects other than A or B, that is, healthy infants, those with only defects B, and those with A plus B ($B_1 + O_1$).

a = Number of infants with defect B in children who also have any other defect (B_2), excluding infants with only defect A plus B (O_1).

b = Number of infants without defect B, but with any other defect, including A in an MCA pattern but excluding cases with only defect A as an isolated anomaly ($D - B - A_1$).

c = Number of cases with defect B but without any other defects except for A. Thus, this group would include only children with defect B and A + B ($B_1 + O_1$).

d = (If we have the total births), the number of infants without defect B and without other defects different from isolated A. Thus, this is constituted by the total nonmalformed infants and those with only defect A as an isolated anomaly ($T - D + A_1$).

[a'], [b'], [c'], [d'] = Number of cases with defect A within each cell.

The ratios

$$a'/a,\ b'/b,\ c'/c,\ \text{and}\ d'/d$$

represent the *specific risks* and can be further explained as follows:

$a'/a = O_2/(B_2 - O_1) = P(A|B \text{ plus other defects})$.

This is, the proportion of infants with defect A among infants with defect B in combination with any other congenital defect.

$b'/b = (A_2 - O)/(D - B - A_1) = P(A|\text{other defects})$.

This is, the frequency of infants with defect A among infants with defects other than B.

$c'/c = O_1/(B_1 + O_1) = P(A|\text{only B})$

This is, the frequency of infants with defect A among the children with defect B and no other defects.

$d'/d = A_1/(T - D + A_1) = P(A|\text{nonmalformed infants})$

This is, the frequency of infants with only defect A as an isolated anomaly in the group of nonmalformed children.

Although the method we propose here is based on a total population of births, it can also be used when only the population of malformed infants is available by covering only cells a, b, and c in the 2 × 2 table.

The quotients between the specific risks give the ralative risk (RR). Thus, the analysis is based in the following RR's:

(1) The times defect A is more frequent in children with defect B (without defects other than A) than in infants with defect C (defects other than B), is given by the following RR:

$$\frac{P(A|\text{only B})}{P(A|\text{others})} = \frac{c'/c}{b'/b} = \frac{O_1}{(B_1 + O_1)} : \frac{(A_2 - O)}{(D - B - A_1)} = RR$$

We can determine if the RR is statistically significant by applying the Fisher exact test to the 2 × 2 table:

$$\frac{O_1}{(B_1 + O_1)} : \frac{(A_2 - O)}{(D - B - A_1)}$$

The quotient between these two ratios indicates whether or not defect A tends to cluster with only defect B more frequently than with defects other than B.

If the tendency of defect A to cluster with defect B is nonspecific, that is, similar to or less than the tendency of defect A to cluster with any other defect (C), the value of the RR should be 1 or less. When the RR has a significant value above 1, there is a *specific* association between defects A and B, provided the infant has no other defect. But, as the analysis also considers the

group of other congenital defects, the obtained RR is influenced by the frequency in which each defect is associated to other defects. That is, if defect A is more frequently associated to any other defect than defect B, its frequency among the group of other defects is higher and, consequently, decreases the value of the relative risk. Thus, the analysis should also be done in the other direction, that is, the proportion of cases with defect B among cases with defect A in relation to the proportion of cases with defect B among other defects (C). This should be studied by replacing A for B and B for A in the last expression to do the analysis in the other direction, as follows:

$$\frac{P(B|\text{only A})}{P(B|\text{others})} = \frac{O_1}{(A_1 + O_1)} : \frac{(B_2 - O)}{(D - A - B_1)} = RR$$

The results of these two analyses will measure the association by considering the tendency of each defect to cluster with other anomalies. If the result in one of the two directions is statistically significantly higher than 1, there is a specific association between A and B when they are the only defects present in the child. When a statistically not significant result is observed in only one of the two directions, it is due to the different proportions in which each defect is associated to congenital anomalies other than A or B, and could indicate the existence of other preferential associations.

(2) The association of two particular defects in infants with or without other congenital anomalies, by the relative risk:

$$\frac{a' + c'}{a + c} : \frac{b'}{b} = \frac{O}{B} : \frac{(A_2 - O)}{(D - B - A_1)} = RR$$

This RR gives the times the proportion of infants with defect A is higher in children with defect B (with or without other congenital anomalies) than in infants with defects other than B. The result of this analysis is similar to that of Khoury et al. [1990] because it analyzes the association of two defects without discriminating whether the infants have only the two studied anomalies or if they also have other congenital defects.

(3) Finally, when the total population of births is available, it is also possible to estimate the number of times in which the proportion of infants with defect A among infants with defect B without defects other than A, is higher than the proportion of infants with only defect A among nonmalformed infants. This information is given by the quotient

$$\frac{P(A|\text{only B})}{P(A|\text{nonmalformed infants})} = \frac{c'/c}{d'/d}$$

$$= \frac{O_1}{(B_1 + O_1)} : \frac{A_1}{(T - D + A_1)} = RR$$

RESULTS

Using the ECEMC data [Martínez-Frías and Urioste, 1994], we selected a group of congenital anomalies to apply this method and compared the results with those observed using the method proposed by Khoury et al. [1990]. The defects were selected because some are included in known sequences, while others do not appear as being pathogenetically related. The selected defects were: neural tube defects (NTD) as a single group, spina bifida (SB), cleft lip with or without cleft palate (CL), anorectal atresia (AN), and tracheoesophageal atresia (TEA). To test the method, we analyzed two groups of defects: (1) the association of SB with club foot (CF), two defects that are known to be part of the SB sequence; and (2) SB plus hypospadias (HY) which, as we previously described [Martínez-Frías, 1994], are preferentially associated and belong to the same midline developmental field defect (DFD).

Table I shows the number of cases, by clinical presentation, with each of the selected defects that will be analyzed. Table II presents the number of infants with each anomaly whether isolated, in combination among themselves only, or with one or more defects (in an MCA pattern including syndromes).

The first part of Table III indicates, as an example, the analysis of the association of TEA and AN, in both directions: the proportion of cases with AN among infants with only TEA in relation to the proportion of cases with AN among infants with other defects, and

TABLE I. Number of Infants With the Defects Studied by Clinical Presentation*

		With MCA patterns		
Defects	Isolated	N	%	Total
Neural tube defects (NTD)	592	137	18.79	729
Cleft lip ± palate (CL)	441	147	25.—	588
Tracheo-esophageal atresia (TEA)	108	94	46.53	202
Anorectal atresias (AN)	90	105	53.85	195
Spina bifida (SB)[a]	354	81	18.62	435
Hypospadias (HY)	1,670	139	7.68	1,809
Club foot (CF)	1,287	957	42.65	2,244

* Total liveborn infants: 1,093,865 (497,294 males); total malformed infants: 21,168 (10,186 males).
[a] Only males: 159 isolated and 36 with MCA patterns.

TABLE II. Number of Infants With Each of the Defects Studied Whether Isolated or in Combination With One or More Defects, and the Total MCA Patterns*

		In MCA patterns										
		Association with										
		CL		TEA		AN		HY		CF		Total with
Defects	Isolated	Only	MCA	Only	MCA	Only	MCA	Only	MCA	Only	MCA	MCA patterns
NTD	592	2	4	0	4	0	8	—	—	—	—	137
CL	441	—	—	0	6	0	4	—	—	—	—	147
TEA	108	0	6	—	—	2	15	—	—	—	—	94
AN	90	0	4	2	15	—	—	—	—	—	—	105
SB	354	—	—	0	3	0	8	6	7	72[a]	33	81
HY	1,670	—	—	—	—	—	—	—	—	—	—	139
CF	1,287	—	—	—	—	—	—	—	—	—	—	957

* Total liveborn: 1,093,865; total malformed infants: 21,168.
[a] These 72 infants with only SB and CF are included in the total of 354 isolated cases of SB. These 72 infants should be included among the infants with MCA patterns only for the analysis of the association of SB + CF and consequently they should be excluded from the isolated cases.

vice versa, the proportion of infants with TEA among the cases with only AN in relation to the proportion of cases with TEA among infants with other defects. The results show that AN is observed among infants with TEA 4.31 times more frequently than among infants with other defects, while TEA is observed among infants with AN 5.89 times more frequently than among infants with other congenital defects. Although these results are not strikingly different, they could be totally different depending on the relationship of each of the analyzed defects with other congenital anomalies. The second part of Table III depicts the times the proportion of infants with AN is higher in infants with TEA with or without other congenital anomalies, than in infants with other defects, and vice versa. The result of this second analysis is similar to that of Khoury et al. [1990] because it analyzes the association of two defects without discriminating whether the children only have the two studied anomalies or whether they also have other congenital defects.

Table IV indicates the results of a two-directional analysis with selected ECEMC defects using our method and that proposed by Khoury et al., [1990]. This table also shows the results observed by Khoury et al. [1990] in Atlanta for some of these combinations. The RR of the first column indicates the times the proportion of one of the two studied defects is higher in infants with the other studied defect (without other congenital

TABLE III. Analysis of the Association of Defects AN and TEA in Both Directions*

Data:
Defect A = AN; defect B = TEA; O = AN + TEA; C = others

Total livebirths: T = 1,093,865
Total malformed children: D = 21,168

Total	Isolated	Multiples
A = 195	$A_1 = 90$	$A_2 = 105$
B = 202	$B_1 = 108$	$B_2 = 94$
O = 17	$O_1 = 2$	$O_2 = 15$

AN among infants with only TEA

$$\frac{c'/c}{b'/b} = \frac{P(A/\text{only B})}{P(A/\text{others})} = \frac{O_1}{(B_1 + O_1)} : \frac{(A_2 - O)}{(D - B - A_1)} = \frac{2}{110} : \frac{88}{20{,}876} = 4.31$$

TEA among infants with only AN

$$\frac{c'/c}{b'/b} = \frac{P(B/\text{only A})}{P(B/C)} = \frac{O_1}{(A_1 + O_1)} : \frac{(B_2 - O)}{(D - A - B_1)} = \frac{2}{92} : \frac{77}{20{,}865} = 5.89$$

AN among TEA plus others defects

$$\frac{O}{B} : \frac{(A_2 - O)}{(D - B - A_1)} = \frac{17}{202} : \frac{88}{20{,}876} = 19.96$$

TEA among AN plus other defects

$$\frac{O}{A} : \frac{(B_2 - O)}{(D - A - B_1)} = \frac{17}{195} : \frac{77}{20{,}865} = 23.62$$

* The numbers are drawn from Tables I and II.

TABLE IV. Comparison of the Results of Our Methods With the Adjusted O/E Proposed by Khoury et al. [1990] and the Observed Results in the Atlanta Population

	Data from the ECEMC							
	As isolated association		In any type of presentation		Using the Khoury et al. method		Khoury et al. [1990] Adjusted O/E	
Type of association	RR[a]	$P<$	RR[b]	$P<$	O/E[c]	$P<$	O/E[c]	$P<$
NTD + CL	0.69	NS[d]	1.56	NS	1.4	NS	4.3	0.05
CL + NTD	0.50	NS	1.17	NS				
NTD + TEA	—		6.52	0.04	2.7	NS	1.4	NS
TEA + NTD	—		1.24	NS				
NTD + AN	—		6.48	0.00006	4.7	0.01	4.7	0.05
AN + NTD	—		2.30	0.03				
CL + TEA	—		4.32	0.004	3.6	0.01	1.8	NS
TEA + CL	—		2.37	0.05				
CL + AN	—		2.95	0.05	1.1	NS	1.2	NS
AN +CL	—		1.38	NS				
AN + TEA	4.31	0.08	19.96	0.000000				
TEA + AN	5.89	0.05	23.62	0.000000	20.—	0.0000	22.1	0.05
TEA + SB	—		1.56	NS	1.—	NS	—	—
SB + TEA	—		3.92	0.05				
SB + AN	—		11.13	0.000001				
AN + SB	—		3.91	0.002	2.9	0.01	—	—
SB + HY	1.28	NS	2.57	0.005	4.6	0.001	—	—
HY + SB	2.40	0.05	4.40	0.00003				
SB + CF	20.58	0.000000	18.17	0.0000000	4.17	0.0000	—	—
CF + SB	4.64	0.000000	5.51	0.0000000				

[a] The times the proportion of one of the two studied defects is higher in infants with only the other studied defect than in children with other congenital anomalies.
[b] The times the proportion of one of the studied defects is higher in children with the other defect (with or without other congenital anomalies) than in infants with other defects.
[c] The times in which the observed association is higher than the expected one, if the association of these defects were not higher than among other defects.
[d] NS, not significant.

defects) than in children with other congenital anomalies. The second column shows the times the proportion of one of the study defects is higher in children with the other defect (with or without other congenital anomalies), than in infants with other defects (C). The result of the analysis in this second column is like the method proposed by Khoury et al. [1990] because it analyzes the association of two defects without discriminating whether the children only have the two anomalies under study or whether they also have other congenital defects. As can be seen in Table IV, all but the NTD + CL associations in this second column are nonrandom. However, for most of the isolated defect combinations, that is, in children with only the two studied defects (first column of Table IV), the associations are not specific and could be associated just by chance. For those defects that belong to the same sequence (such as SB and CF) or to the same DFD (such as AN and TEA or SB and HY), the associations are specific and nonrandom, either as an isolated association (first column) or as part of a more generalized MCA pattern (second column). The results observed in each direction are affected by the proportion in which the infants with each of the analyzed defects are associated to other congenital anomalies. For instance, Table I shows that 18.62% of the cases with SB are observed in children with an MCA pattern, while 42.65% of the cases with CF are part of MCA patterns. This explains why the proportion of CF in infants with SB is smaller than the proportion of SB in infants with CF, and indicates that CF could also be preferentially associated to other defects different from SB.

Table IV also includes the analyses of our data using the Khoury et al. method (third column) as well as the results observed in the Atlanta population (fourth column) by Khoury et al. [1990].

DISCUSSION

Several different methods have been proposed for analyzing the patterns of congenital anomaly clusters [Roberts and Powell, 1972; Källén, 1988; Khoury et al., 1990]. All of them analyze the presence of two or more defects in the same infant without taking into account if the child has or does not have other congenital defects.

Roberts and Powell [1972] suggested that birth defect associations are generalized and nonspecific, since many defects showed large O/E ratios when the syntropy index was used. Källén [1988], based on the same idea, proposed an analysis of defect clustering using the subset of infants with MCA patterns. However, Khoury et al. [1990] showed that the method proposed by Källen [1988] was limited and conservative because the O/E ratio was affected by the tendency of the defects to cluster among themselves. Khoury et al. [1990] proposed a new method for adjusting for this tendency. With their approach the adjusted O/E ratio indicated the times in which the observed number of children

with the combination of two particular defects (in any type of clinical presentation) was higher than the expected one if the association between those defects was not higher than among two others.

Aside from the problem of controlling for the nonspecific tendency of defects to cluster among themselves, one of the main interests in the study of infants with MCA patterns is to elucidate whether the association of *only two* defects in the same child, is the result of a chance association, a nonspecific and generalized tendency for both defects to cluster, or a specific nonrandom association. The method presented here, allows the analysis of the following aspects:

(1) The specific association of *only two* particular defects in infants who *do not have* other primary congenital anomalies by controlling the nonspecific tendency of the defects to cluster among themselves. This is estimated through a RR which reflects the times the proportion of infants with defect A among infants with *defect B without defects other than A* is higher than the proportion of infants with defect A among infants with defects other than B and vice versa. Thus, this RR gives the value of the specific association between defects A and B in infants without other congenital defects. Nevertheless, since the frequency in which each defect is associated to other defects could influence the results, the analysis should be done in the other direction, at least if the first result is not statistically significant. On the other hand, the differences between the two-direction analysis may give us clues on other possible associations.

(2) The association of two particular defects in infants with or without other congenital anomalies. This is similar to the analysis of Khoury et al. [1990] because it does not discriminate whether the studied association is in infants with only the two studied defects.

(3) When the total population of births is available, it is also possible to estimate the number of times in which the proportion of infants with defect A among infants with defect B with or without defects other than A, is higher than the proportion of infants with only defect A among nonmalformed infants.

The presence of only SB and CF (as an isolated association) in the same child was analyzed as a test of the proposed methods, since these defects are part of the same sequence and, consequently, should be preferentially associated. As expected, the results show that SB and CF constitute a nonrandom association, whether they are the only defects present in the child or whether they are part of a more extensive MCA pattern. The same occurs with those defects that are part of the same DFD, such as SB and HY, which, as was previously observed [Martínez-Frías, 1994], are preferentially associated and belong to the midline DFD. Thus, the associations between SB and HY as well as between TEA and AN (two blastogenetic midline defects) in children who do not have other defects, are specific since the result is statistically significant in one of the analyses. The absence of significance in the other direction could well be due to the fact that SB and AN are associated to other defects more frequently than HY and TEA.

In general, our results show that in most of the cases, the presence of only two defects in the same child is more probably due to chance, except for those defects that may have a common pathogenesis such as those that are part of the same sequence or constitute a DFD.

As a comparison, Table IV presents the O/E ratios derived from the analysis of our data using the method proposed by Khoury et al. [1990]. As can be seen, the O/E ratios were small and several of them lost their statistical significance. The differences observed analyzing our cases using the method proposed by Khoury et al. [1990], as compared with the results of our method, could well be due to their use of all malformed infants. Nevertheless, the results of Khoury et al. [1990] are closer to ours in the analysis of the proportion of infants with defect A among those with defect B *with or without* other congenital defects in relation to the proportion of infants with defect A among those with defects other than B (second column of Table IV). Thus, the difference between the method proposed by these authors and ours, is that with our method we can discriminate if the association of *only two* defects is specific or not.

Table IV also includes the results observed by Khoury et al. [1990] in the Atlanta population. All the differences with our results in the second column, except for the association of NTD + TEA, are related to the associations which include cleft lip with or without cleft palate. We suspect that the ethnic-genetic differences between the Atlanta population and the Spanish one would account for these results, since the black population has a different frequency of oral clefts than the Spanish population.

In conclusion, we present a method to analyze the specific association of defined congenital defects in the same child which, like the method proposed by Khoury et al. [1990], takes into consideration the generalized and nonspecific tendency for particular defects to cluster with any defect. Our method allows an examination not only of the association between two selected defects in infants with or without other congenital anomalies, but also the association of *only two* defects of interest in children who do not have any other anomaly. This is important because it may give us clues to possible nonrandom patterns of defects even when they are constituted by only two anomalies, and help to elucidate whether they constitute a chance association, or represent the generalized tendency of two defects to cluster among themselves. Additionally, it might also indicate their common pathogenesis (such as identifying DFD or sequences). Finally, the analysis of the association of only two congenital defects in two directions, may give clues on other possible preferential associations.

ACKNOWLEDGMENTS

This work was supported in part by a grant from Dirección General de Salud Pública, Ministerio de Sanidad y Consumo of Spain, and by Fundación 1000 para la investigación sobre defectos congénitos of Spain.

REFERENCES

Källén B (1988): "Epidemiology of Human Reproduction." Boca Raton, Florida: CRC Press Inc, pp 31–35.

Khoury MJ, James LM, Erickson JD (1990): On the measurement and interpretation of birth defect associations in epidemiologic studies. Am J Med Genet 37:229–236.

Martínez-Frías ML (1994): Spina bifida and hypospadias: A non-random association or an X-linked recessive condition? Am J Med Genet 52:5–8.

Martínez-Frías ML, Urioste M (1994): Segmentation anomalies of the vertebras and ribs: A developmental field defect: Epidemiologic evidence. Am J Med Genet 49:36–44.

Roberts CJ, Powell RG (1972): Interrelation of the common congenital malformations: Some etiological implications. Lancet 2:848–850.